

コーパスに基づく発話意図タイプ決定ルールの自動生成

熊本 忠彦[†] 伊藤 昭^{††}

筆者らは、話し言葉対話システムの自然言語処理部として発話文（漢字かな混じりのテキスト文）から発話意図記述（意味表現）を生成する手法を提案している。発話意図記述は、一種の発話内行為タイプである発話意図タイプと発話文の命題内容（もの、行為、状態に関する記述）を表現するフレームから成る。発話意図タイプは、支援対話における支援要請の種類という観点から12タイプに分類されている。本手法において、発話意図タイプの決定は文法的知識に基づいて人手で設計されたヒューリスティックなルールを用いて行われた。そこで、本論文では、発話意図タイプ付発話文コーパスを作成し、そのコーパスから発話意図タイプ決定ルールを自動生成するためのアルゴリズムを提案する。また、提案アルゴリズムの性能を評価するために、本アルゴリズムによって生成されたルールと一般的な学習手法であるC4.5アルゴリズムによって生成されたルールの性能を比較し、その有効性を検討する。その結果、提案アルゴリズムによって作られたルールの true error rate（open データに対する error rate）が最も小さく、精度において良い結果を示した。

Automatic Generation of Utterance Type Determination Rules from an Annotated Corpus

TADAHIKO KUMAMOTO[†] and AKIRA ITO^{††}

We have proposed as the natural language understanding module of our dialogue system the method which constructs a semantic structure, called a communicative intention (CI) description, from a spoken sentence. A CI description consists of a CI type and an utterance proposition. The latter represents "objects," "actions," and "states" that can be defined in a task domain, and is described in frame representation, while the former corresponds to an illocutionary force type in terms of the speech act theory, and there are 12 CI types. However, the type determination rules which the method for constructing a CI description used in order to determine the CI type of a spoken sentence were hand-made. In this paper we construct an annotated corpus consisting of spoken sentences with CI types, and propose a method for generating type determination rules from the annotated corpus.

1. はじめに

筆者らは、話し言葉による対話を通じてユーザの計算機利用を支援するシステム（支援システム）^{1)~4)}を構築しており、その自然言語理解部としてユーザ発話文（漢字かな混じりのテキスト文）から発話意図記述（意味表現）を生成する手法²⁾を提案している。発話意図記述は、発話文の命題内容（もの、行為、状態に関する記述）を表現するフレームと一種の発話内行為タイプである発話意図タイプから成る。この発話意図タイプは、支援対話コーパス⁵⁾の分析に基づいて設計さ

れたものであり、支援要請の種類という観点から12タイプに分類されている。本手法において、発話意図タイプの決定は文法的知識に基づいて設計されたヒューリスティックなルールを用いて行われた。

近年、コーパスを利用した研究がさかに行われており、言語知識や共起知識、翻訳ルールといった様々なヒューリスティクスを獲得するための手法^{6)~9)}が提案されている。特に、文献9)では発話文（音声認識結果に相当）を意味表現に直接変換するための言語モデルをコーパスから自動的に獲得する手法が提案されており、コーパスに基づく処理の新たな可能性が示されている。一方、いわゆる発話内行為タイプに関する研究も多く、対話管理や次発話予測などに利用されている^{10)~12)}。しかしながら、発話文へのタイプ付与を自動的に行う研究は少なく、多くの場合、人手で行われている^{10)~14)}。そのため、労力を要するうえ、一貫

[†] 郵政省通信総合研究所関西支所
Kansai Advanced Research Center, Communications
Research Laboratory, MPT

^{††} 山形大学工学部
Faculty of Engineering, Yamagata University

性に乏しかった。また、自動的に行う場合^{2),15)~18)}でも、手法の核となる部分は人手で設計されたヒューリスティックなものであり、研究を進めるうえで重要となる大規模なタイプ付コーパスを作成することが難しい。たとえば、文献15)では3種類の情報(韻律、単語列、談話文法)を用いてタイプ付与を自動的に行う試みが示されており、各情報の利用方法や統合の仕方が提案されているが、当研究は大規模コーパスへのタイプ付与を意図したものではなかった。

そこで本論文では、発話意図タイプ付発話文コーパスから発話意図タイプ決定ルールを自動生成するアルゴリズムを提案する。また、提案アルゴリズムの性能を評価するために、本アルゴリズムによって生成されたルールおよび文献2)のルール、一般的な学習手法であるC4.5アルゴリズムによって生成されたルールの性能を比較し、その有効性を検討する。

なお、支援システムがサポートするタスク領域として、UNIX 計算機上のXウィンドウ下で動作する電子メール処理プログラムXMH¹⁰⁾が設定されている。XMHは、電子メールを受け取る/表示する/作成する/送る/保管するといった機能を有し、主にマウスで操作される。

以下、本論文の構成について述べる。2章ではルール生成用データおよび評価実験用データとして利用される発話意図タイプ付発話文コーパスについて述べる。3章では文献2)で提案された発話意図タイプ決定手法の枠組みを示す。4章ではタイプ付コーパスから発話意図タイプ決定ルールを生成するためのアルゴリズムを提案する。5章では提案アルゴリズムの性能をtrue error rate(評価実験用のopenデータに対するerror rate)を用いて評価する。最後に6章では、本論文のまとめと今後の課題について述べる。

2. 発話意図タイプ付発話文コーパスの作成

筆者らは、支援システムの代わりに人間コンサルタントが支援するという「対話による支援」実験を行い、ユーザの自由かつ自然な発話を収集した。そして、収集されたユーザ発話を書き起こすことによって、支援対話データベース(ユーザ42人分の発話文855文を含むテキストデータベース)を作成した⁵⁾。

この支援対話データベースから抽出されたユーザ発話文(フラグメント文24文を除くユーザ16人分の発話文475文)に発話意図タイプを付与し、発話意図タイプ付発話文コーパスを作成した。このタイプ付与は、その一貫性を保つために筆者らのうちの1人によって行われた。なお、自然言語処理研究者2人にこの発話

表1 発話意図タイプと支援要請の種類
Table 1 Classification of utterance types.

タイプ	支援要請の種類
手法	手法に関する質問
概念	概念に関する質問
属性値	属性値に関する質問
属性値:OK	事象そのもの 未来事象の正当性
真偽値	真偽値に関する質問
真偽値:OK	事象そのもの 未来事象の正当性
真偽値:NO	未来事象の禁止性
真偽値:EQ	その他
信念	事実等の表明
ゴール	タスクゴールの表明
対話開始	対話開始の合図
対話終了	対話終了の合図

意図タイプ付与を行ってもらったところ、コーパスとの一致率は平均で81.9%であった。また、ユーザ発話文には、日常会話文と同様、間投表現や言い直し、言い淀み、言い換え表現が現れていたが、これらはすべてコーパス作成にあたって取り除かれた。このような表現は、タイプ決定より前の処理段階で除去/修正されるべきものであり、提案アルゴリズムや比較対象としているアルゴリズムへの影響はないものと考えられる。一方、助詞落ちや助詞の誤運用、倒置なども観測されたが、これらはそのまま残されている。

なお、発話意図タイプには、ユーザ発話に対する応答生成パターンを選択するためのマーカーとしての役割が与えられており、支援要請の種類と表1に示されたような対応関係にある。表1において、未来事象の正当性に関する質問とは「～したらいいんですか」のようなある事象の容認可能性を問うことによって許可を求める文のことであり、未来事象の禁止性に関する質問とは、「～してはいけないんですか」のような禁止の形式で問うことによって許可を求める文のことであり、また、事象そのものに関する質問とはXMH上で生じた事象(出来事/状態)に関する質問のことであり、その他のことに関する質問とはXMH上のある状態の正当性に関する質問のことであり。各発話意図タイプに対応するユーザ発話文の例を表2に示す。

3. 発話意図タイプ決定手法の枠組み

本章では文献2)で提案された発話意図タイプ決定手法の枠組みを示す。

3.1 発話意図タイプ決定用特徴量の定義

発話意図タイプを決定する際に用いられる特徴量には、意味語、品詞情報(品詞名/品詞細分類名)、活用

表2 発話意図タイプとユーザ発話文の例

Table 2 Examples of sentences spoken by users.

タイプ	ユーザ発話文
手法	移すのはどうするんですか
概念	スクロールって何ですか
属性値	住所忘れました
属性値:OK	何を書いたらいいんですか
真偽値	これで移動できてるんですか
真偽値:OK	終了を押せばいいんですね
真偽値:NO	これ押さなくちゃ駄目なんですか
真偽値:EQ	メッセージはこれだけでいいんですか
信念	移動できてないみたいなんですけど
ゴール	ロイの次にカーソルを持っていきたい
対話開始	すみません
対話終了	わかりました

表3 意味語の一覧

Table 3 A list of semantic words.

意味語	形態素の例
対話開始合図	すみません, あの
対話終了合図	わかりました
HOW 句	どうする, どうやる, どう
教示述語	教える
受益述語	もらう, くれる
欲求述語	欲しい
疑問語	何処(「何」を除く疑問詞)
何	何
OK 述語	いい, よい, よろしい
欠如述語	忘れる, わからない
禁止述語	駄目だ, いけない
疑問助詞	か, ね, よね(終助詞)
文末	(文末)
手法語	仕方, 方
する	する
何でも	何でも(陳述副詞)
動態述語	読む, 消す(動態動詞)
状態述語	ある, 違う(状態動詞) おかしい, ない(形容詞) 読める, 消せる(可能動詞) 消える, 出る(無意志動詞)
知識述語	知る, 覚える, わかる

形情報, 原形情報の4種類がある。意味語は, 本論文では発話意図タイプの決定に影響を与える意味情報の種類と定義され, 文献20)の記述を基に設計された。この意味語は入力形態素の原形情報と品詞情報から辞書引きされる。たとえば, 話者の「良い」という判断を表す形容詞「いい, よい, よろしい, …」からは意味語「OK 述語」が抽出され, 疑問を表す終助詞「か, ね, よね, …」からは意味語「疑問助詞」が抽出される。ここで, 意味語の一覧を表3に示す。

一方, 名詞や指示詞, 判定詞からは品詞情報(サ変名詞, 普通名詞, 形式名詞, 副詞的名詞, 名詞形態指

表4 接尾辞等による特徴量の交換規則

Table 4 Rules for substituting feature values.

直前の特徴量	入力形態素	交換後の特徴量
“サ変名詞”	“する”	“動態述語” + (“する”の活用形)
“サ変名詞”	“できる”	“状態述語”
“知識述語”	“ない”	“欠如述語”
(活用形)	“ます”	(“ます”の活用形)
(活用形)	“せる”	(“せる”の活用形)
“動態述語” + (活用形)	“たい”	“欲求述語”
“動態述語” + “タ系連用テ形”	“ます”	“状態述語”
“動態述語” + (活用形)	“ない”	“状態述語”
“動態述語” + (活用形)	“いる”	“状態述語”
“動態述語” + (活用形)	“れる”	“状態述語”
“動態述語” + (活用形)	“る”	“状態述語”

示詞, 判定詞)が抽出され, 動態動詞からは活用形情報が抽出される。また, 提題助詞(「とは」, 「って」, 「は」など)や格助詞の「で」, 取り立て助詞の「も」, 「でも」からは原形情報(とは, って, は, で, も, でも)が抽出される。また, 動詞「する」や「できる」, 動詞性接尾辞「れる」や「いる」, 「る」, 「せる」, 「ます」, 形容詞性述語接尾辞「ない」や「たい」は, 直前に抽出された特徴量を別の特徴量に変換する作用がある。この交換規則を表4にまとめる。

3.2 発話意図タイプ決定手法

形態素解析および特徴抽出:

入力発話文は, まず汎用日本語形態素解析システムJUMAN²¹⁾によって形態素解析される。JUMANは, 入力発話文を適当な形態素に分解し, 各形態素に原形情報や品詞情報, 活用形情報などを付加する。なお, 本論文の文法用語は, JUMANで採用されている益岡・田窪文法²⁰⁾に依っている。

発話意図タイプ決定用特徴量は, JUMANの出力である形態素列から抽出される。たとえば, 図1に示されるように発話文「終了を押せばいいんですね」からは特徴リスト「(サ変名詞 動態述語 基本条件形 OK 述語 疑問助詞 文末)」が得られる。

発話意図タイプの決定:

入力発話文の発話意図タイプは, その発話文から抽出された特徴リストに発話意図タイプ決定ルールを適用することによって決められる。ここで, 文献2)で得られたルールを例として図2に示す。なお, ルールの記述形式は以下のとおりである(拡張バックス記法)。

形態素	特徴量
終了	サ変名詞
を	
押せば	動態述語 基本条件形
いい	OK 述語
んです	
ね	疑問助詞
	文末

図1 発話意図タイプ決定用特徴量の抽出

Fig. 1 Morphological analysis and feature extraction.

rule-1:	((対話開始合図))	対話開始
rule-2:	((対話終了合図))	対話終了
rule-3:	((HOW 句))	手法
rule-4:	((教示述語))	属性値
rule-5:	((受益述語))	ゴール
rule-6:	((欲求述語))	ゴール
rule-7:	((疑問語)(OK 述語))	属性値:OK
rule-8:	((疑問語)(基本形 文末))	属性値:OK
rule-9:	((疑問語))	属性値
rule-10:	((何 も)(OK 述語))	真偽値:OK
rule-11:	((何 も))	信念
rule-12:	((とは 何))	概念
rule-13:	((って 何))	概念
rule-14:	((何 判定詞))	手法
rule-15:	((何 する))	手法
rule-16:	((何))	属性値:OK
rule-17:	((状態述語 OK 述語))	真偽値:OK
rule-18:	((動態述語)(OK 述語))	真偽値:OK
rule-19:	((OK 述語))	真偽値:EQ
rule-20:	((普通名詞 欠如述語))	属性値
rule-21:	((状態述語 欠如述語))	属性値
rule-22:	((サ変名詞 欠如述語))	概念
rule-23:	((疑問助詞 欠如述語))	真偽値
rule-24:	((欠如述語))	手法
rule-25:	((禁止述語))	真偽値:NO
rule-26:	((基本形 疑問助詞))	真偽値:OK
rule-27:	((疑問助詞))	真偽値
rule-28:	((基本形 文末))	真偽値:OK
rule-29:	((文末))	信念

図2 文献2)で得られた発話意図タイプ決定ルール

Fig. 2 Type determination rules obtained in Ref. 2).

〈ルール〉 ::= (〈パターン部〉 〈発話意図タイプ〉)
 〈パターン部〉 ::= (〈パターン〉) | (〈パターン〉
 〈パターン〉)
 〈パターン〉 ::= (〈特徴量〉+)
 〈特徴量〉 ::= 〈意味語〉 | 〈品詞情報〉 | 〈活用形
 情報〉 | 〈原形情報〉

ルールの適用は次のようにして行われる。

【ルール r ($1 \leq r \leq R$) の適用】

入力特徴リストにルール 1 からルール R を順に適用する。入力特徴リストとルール r のパターン部とのマッチングに成功したら、対応する発話意図タイプを出力とし、ルールの適用を終了する。

【パターン部のマッチング】

パターン部の最も左にあるパターンから順にマッチングを行い、全パターンのマッチングに成功したら、そのパターン部のマッチングは成功である。

【パターンのマッチング】

当該パターンが入力特徴リスト中のある部分列と同一であるならば、そのパターンのマッチングは成功である。成功したら、その部分列よりも文末方向にある特徴量の列を新たに入力特徴リストと見なし、次のパターンとのマッチングを行う。

たとえば、図1で得られた特徴リスト「(サ変名詞 動態述語 基本条件形 OK 述語 疑問助詞 文末)」に図2のルールを適用してみると、ルール18のパターン部「((動態述語)(OK 述語))」とのマッチングに成功し、その結果、「真偽値:OK」が出力される。

4. N グラムを用いた発話意図タイプ決定ルールの自動生成

本章では、事例(特徴リストと発話意図タイプとの組)から発話意図タイプ決定ルールを自動生成するためのアルゴリズムについて述べる。本アルゴリズムへの入力事例の集合であり、出力はルールの順序付集合となる。

提案アルゴリズム(意味語と発話意図タイプを利用した部分空間探索法)

(1) 優先意味語の決定

(a) 意味語に基づく部分集合の作成

各意味語ごとに、その意味語を含む事例を入力から抽出し、部分集合を作る。ただし、1個の事例が複数の部分集合に重複して含まれてもよい。なお、意味語は表3に示されるように19種類あるので、19個の部分集合が作られる。

(b) 発話意図タイプの異なり数と事例数の計数

各部分集合ごとに、観測される発話意図タイプの異なり数と各発話意図タイプごとの事例数を求める。

(c) 優先意味語の選定

該当する部分集合が1つになるまで、以下の条件を順次適用する。

条件1 事例数が閾値(2個)以上である発話意図タイプを有する。ただし、条件に該当する部分集合が1つもない場合は、閾値を1にする。

条件2 発話意図タイプの異なり数が最も少ない。

条件3 事例数が最も多い発話意図タイプを有する。

条件4 総事例数が最も多い。

条件5 先に得られた。ただし、意味語「文末」に基づく部分集合があれば、その部分集合を優先する。

以上の結果選ばれた部分集合に対応する意味語が優先意味語となる。

(2) 発話意図タイプの順序付け

操作 (1) で選ばれた部分集合において、事例数の少ない発話意図タイプから順に優先順位を付ける (この優先順位の付け方の意味は 5 章で述べる)。以下、優先権のある発話意図タイプを優先発話意図タイプと呼ぶ。

(3) N グラムの作成

入力中の各特徴リストから優先意味語を含む N グラムを作成する。たとえば、優先意味語が「動態述語」のとき特徴リスト「(動態述語 基本形 文末)」からは以下のような N グラム (uni-gram, bi-gram, tri-gram 各 1 個) が得られる。

(動態述語) (動態述語 基本形) (動態述語 基本形 文末)

(4) 頻度の計数

各異なり N グラムに対し、各発話意図タイプごとの事例数を求める。

(5) 最適 N グラムの抽出

該当する異なり N グラムが 1 つになるまで、以下の条件を順次適用する。

条件 1 優先発話意図タイプ的事例数が閾値 (2 個) 以上であり、優先発話意図タイプよりも多い事例を持つ発話意図タイプがない。ただし、該当する異なり N グラムがない場合は、閾値を 1 にする。それでもない場合は、優先発話意図タイプを一時的に次の優先順位のものにする。

条件 2 発話意図タイプの異なり数が最も少ない。

条件 3 事例数が最も多い優先発話意図タイプを有する。

条件 4 N の値が最も小さい。

条件 5 先に得られた。

(6) 発話意図タイプ決定ルールの登録

操作 (5) で選ばれた異なり N グラムと優先発話意図タイプとの組合せをルール r として登録する。このとき、異なり N グラムがそのままパターン部となり、優先発話意図タイプがルール適用に成功したときの出力となる。

(7) 終了条件のテスト

生成されたルールとのマッチングに成功する事例を入力から削除する。残った入力中に優先意味語を含む事例があれば操作 (3) に進み、なければ操作 (1) に進む。ただし、優先意味語を含む事例の中に優先発話意図タイプを含むものがなければ、次の優先順位を持つ発話意図タイプを優先発話意図タイプとする。入力が無くなったなら、操作 (8) を行った後、実行を終了する。

(8) ルールの洗練

まず、最後に生成されたルール R のパターン部を「((文末))」にする。特徴量「文末」はすべての特徴リストに現れるため、ルール R はルール $1 \sim R-1$ とマッチングしなかったすべての事例とマッチングする。次に、 $r = R, \dots, 2, 1$ の順に「ルール $r-1$ とルール r の発話意図タイプが同じであり、かつルール $r-1$ にマッチングする特徴リストがすべてルール r にもマッチングするとき、ルール $r-1$ を削除する」という操作を行う。この結果が最終的な出力となる。

5. 提案アルゴリズムの性能評価

本章では提案アルゴリズムの性能を評価するために、本アルゴリズムによって生成されたルール、および文献 2) で述べたルール、一般的な学習手法である C4.5 アルゴリズム²²⁾によって生成されたルールの性能を比較し、その有効性を検討する。

まずはじめに、文献 2) で述べた条件と同じ条件で評価実験を行った。すなわち、2 章で述べた発話意図タイプ付発話文コーパス中のデータ (475 個) をルール生成用データ (298 個) と評価実験用データ (177 個) に分け、ルール生成用データに対し形態素解析と特徴抽出を行った。形態素解析は、ルール生成用の 298 文中 47 カ所、評価実験用の 177 文中 53 カ所で失敗したが、特徴リストの抽出が正しく行われなかった発話文は、ルール生成用、評価実験用ともに 6 文のみであった。この 12 文については特徴リストを修正した。以上の結果得られた特徴リストと発話意図タイプのペア集合に対し、各アルゴリズムを用いて発話意図タイプ決定ルールの生成を行った。

[提案アルゴリズムによるルール生成]

4 章で示したアルゴリズムによりルールを生成した。生成されたルールの一部を図 3 に示す。

[文献 2) の場合]

文献 2) で作成されたルール (図 2) は、上述のルール生成用データを文法的知識に基づいて整理した結果得られたものであり、今回の評価実験にもそのまま利用できる。

[C4.5 アルゴリズムによるルール生成]

(1) N グラムの作成

入力中の各特徴リストから N グラムを作成する。ただし、N の最大値を 2 とした。したがって、たとえば、特徴リスト「(動態述語 基本形 文末)」からは以下のような N グラム (3 つの uni-gram, 2 つの bi-gram) が得られる。

(動態述語) (基本形) (文末)

((HOW 句))	手法
((対話開始合図))	対話開始
((対話終了合図))	対話終了
((欲求述語))	ゴール
((普通名詞 は 疑問語))	属性値:OK
((疑問語 で))	属性値:OK
((疑問語 普通名詞 動態述語 タ系連用テ形))	属性値:OK
((疑問語))	属性値
((欠如述語 禁止述語))	属性値
((禁止述語))	真偽値:NO
((手法語 は))	真偽値
((手法語))	手法
((って する))	真偽値:OK
((です る))	真偽値
((する))	手法
((何 動態述語 基本条件形 OK 述語))	属性値:OK
⋮	⋮
((サ変名詞 欠如述語))	概念
((欠如述語))	属性値
((基本形 文末))	真偽値:OK
((文末))	信念

図3 提案アルゴリズムによって生成された発話意図タイプ決定ルール (一部)

Fig. 3 Type determination rules generated by the proposed algorithm.

(動態述語 基本形) (基本形 文末)

なお、Nの最大値は、1~7の範囲でルール生成を行い、下述の方法で性能評価した結果に基づいて決められた値である。

(2) C4.5 アルゴリズム用データの作成

C4.5 アルゴリズムは、決定木学習アルゴリズムなので、データ (属性・属性値の集合とクラス) の集合を入力とし、決定木を出力とする。そこで、入力中の各ペアにおいて、特徴リストから作成されたNグラム1つ1つを属性、発話意図タイプをクラスとした。また、属性値は「真」または「偽」の2値とし、各属性には属性値「真」を与えた。なお、各ペアにおいて、そのペアに存在しない属性 (Nグラム) に対しては、属性値「偽」を便宜上与えた。

(3) C4.5 アルゴリズムの実行

(2) で得られたデータ集合に対して、C4.5 アルゴリズムを適用した。その結果、20 ノードからなる決定木が生成された。この決定木を図4に示す。各ノードはNグラム形式 (uni-gram または bi-gram) で記述されており、そのテスト値は、対応するNグラムが特徴リストの部分集合であれば「真」、なければ「偽」となる。

以上の方法により生成された各ルールの true error rate (評価実験用の open データに対する error rate)

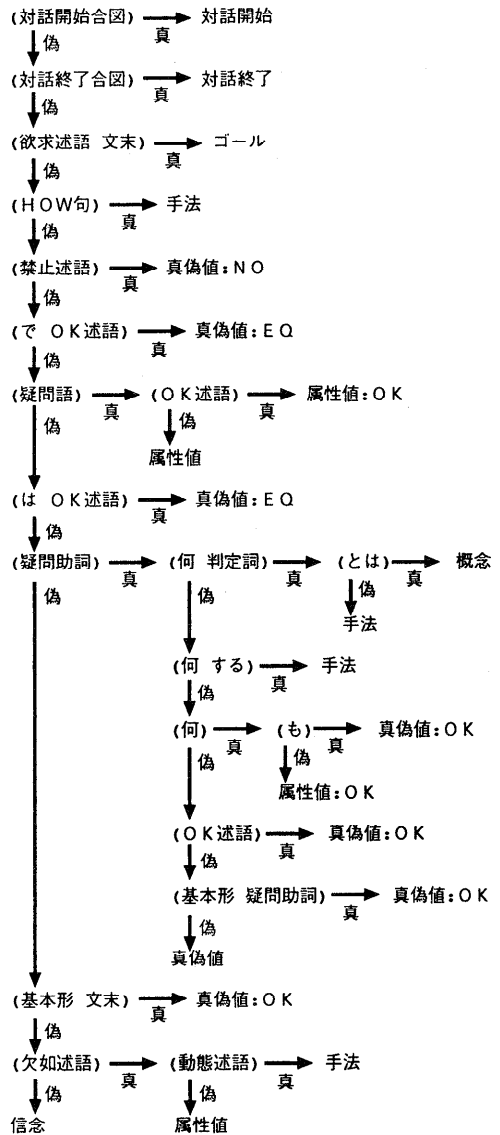


図4 C4.5 アルゴリズムによって生成された発話意図タイプ決定用決定木

Fig. 4 Decision tree generated by C4.5 algorithm.

表5 True error rate とルール数
Table 5 True error rate and No. of rules.

ルール生成法	true error rate	ルール数
提案アルゴリズム	0.090	39
文献2)の場合	0.107	29
C4.5 アルゴリズム	0.136	37

とルール数を表5に示す。ただし、C4.5 アルゴリズムの項のルール数は、生成された決定木を3.2節で述べたルール記述形式で書き直した場合の数値である。なお、各発話意図タイプの出現率の2乗の和として定義される偶然の正解率を計算してみたところ、0.169

という値であった。

文献2)のルールは、人手によって作成されたため、ルール数29とコンパクトなものとなっている。また、true error rateも0.107と比較的小さい。C4.5アルゴリズムによって作られたルールは、ノード(条件部)の数は20個と少なかったが、ルール数に換算すると37であり、true error rateは若干高く0.136であった。提案アルゴリズムによって作られたルールは、true error rateが最も小さく、精度において良い結果を示した。このアルゴリズムは、意味語を用いて探索空間を限定し、さらにその探索空間の中で出現頻度の低い発話意図タイプ用のルールを優先的に生成する。これは、言い換えれば、個別対応的なルールを優先的に生成することで、より汎用的なルールの生成を可能にしているといえる。このことを実証するために、提案アルゴリズムによって生成されたルールに対し、以下のように定義されるルール占有率とルール適用率を求めた。

$$\text{ルール占有率} = \frac{\text{該当する頻度を持つルールの数}}{\text{総ルール数}}$$

$$\text{ルール適用率} = \frac{\text{該当する頻度を持つルールの生成に寄与したデータの数}}{\text{総データ数}}$$

ただし、ルール占有率とルール適用率は、生成されたルールが個別対応的なルール(頻度1)か汎用的なルール(頻度10以上)かを示すための指標であり、前者はルール全体の中での個別対応的なルール/汎用的なルールの割合を示し、後者はルール生成用データ全体の中での個別対応的なルール/汎用的なルールの生成に寄与したデータの割合を示している。以上の結果を表6に示す。なお、括弧内の数字は該当する頻度を持つルールの数または該当する頻度を持つルールの生成に寄与したデータの数を示している。

ここで、提案アルゴリズムとC4.5アルゴリズムの性能をより客観的に調べるために、5-fold cross-validation²³⁾に基づく評価実験を行った。1回の実験で用いられたルール生成用データは380個、評価実験用データは95個であった。各実験において生成されたルールの平均true error rateを表7に示す。表5では提案アルゴリズムの方が高精度であったが、表7ではほぼ同じ精度である。この結果の違いはルール生成用データの数と評価実験用データの数の割合に依存しているものと考えられ、ルール生成用のデータが比較的少ないとき、提案アルゴリズムの方が有効と考えられる。このことを実証するために、ルール生成用デー

表6 ルール占有率とルール適用率

Table 6 Ratio of No. of rules and ratio of No. of data matched with their rules.

頻度	ルール占有率	ルール適用率
10~	0.20 (9)	0.81 (241)
2~9	0.27 (12)	0.11 (34)
1	0.52 (23)	0.08 (23)
合計	1 (44)	1 (298)

表7 5-fold cross-validationに基づく性能評価
(学習データが多いとき)

Table 7 Evaluation of performance for a large number of learned data by 5-fold cross-validation.

	平均 true error rate
提案アルゴリズム	0.074
C4.5アルゴリズム	0.078

表8 5-fold cross-validationに基づく性能評価
(学習データが少ないとき)

Table 8 Evaluation of performance for a small number of learned data by 5-fold cross-validation.

	平均 true error rate
提案アルゴリズム	0.174
C4.5アルゴリズム	0.184

タを95個、評価実験用データを380個にして、5-fold cross-validationに基づく評価実験を行ってみた。その結果、表8のような結果が得られ、学習データが少ないときの提案アルゴリズムの有効性が示された。

また、それぞれの方法で生成されたルール(図2、図3、図4参照)を見てみると、ルールの条件部において意味語が重要な役割を果たしていることが分かる。意味語は、対話のドメインがソフトウェアの利用支援であることを前提に文献20)の記述に基づいて設計されたドメイン依存の辞書であり、入力形態素の原形情報と品詞情報を用いて辞書引きされる。したがって、条件部に現れる意味語を原形情報と品詞情報の組として置き換えることも可能ではあるが、提案アルゴリズムの性能向上に果たしている役割も大きい。そこで、意味語の設計が定量的に見て全体的な処理のどの程度を占めているのかということ調べるために、次のような実験を行った。すなわち、Nグラム代わりに、意味語のunigramを用いて、提案アルゴリズムを適用し、5-fold cross validationに基づく評価実験を行ってみた。その結果、意味語のunigramからなるルールの平均true error rateは、ルール生成用データ380文、評価実験用データ95文の場合で0.259、ルール生成用データ95文、評価実験用データ380文の場合で0.319という結果であった。表7や表8に示されてい

る結果と比べて高い値となっていることが分かる。

なお、C4.5 アルゴリズムによるルール生成では入力からルールを生成するのに1秒程度しかかからなかったが、提案アルゴリズムには高速化のための工夫がほとんど施されていないため、1分程度を必要とした。ルール生成の高速化を今後の課題とする。

6. ま と め

本論文では、ソフトウェア利用支援ドメインにおいて作成された発話意図タイプ付発話文コーパスから発話意図タイプ決定ルールを自動生成するためのアルゴリズムを提案した。1つのルールは、パターン部と発話意図タイプからなり、パターン部はNグラム（特徴量の列）の形をしている。

本論文では、提案アルゴリズムの有効性を示すために、本アルゴリズムによって生成されたルールと文献2)のルール、C4.5アルゴリズムによって生成されたルールの性能を比較した。その結果、提案アルゴリズムによって生成されたルールの true error rate (open データに対する error rate) が最も小さく、精度において良い結果を示した。提案アルゴリズムは、意味語を用いて探索空間を限定し、さらにその探索空間の中で出現頻度の低い発話意図タイプ用のルールを優先的に生成する。これは、言い換えれば、個別対応的なルールを優先的に生成することで、より汎用的なルールの生成を可能にしているといえる。

今後の課題として、ルール生成アルゴリズムの高速化や離散的なNグラムを用いたルール生成といったことを行いたい。また、文献2)で設計された発話意図タイプや発話意図タイプ決定法の他ドメインへの拡張性についても検討したい。

参 考 文 献

- 1) 熊本忠彦, 伊藤 昭, 海老名毅: 電子メール利用支援システム, 通信総研第 87 回研究発表会予稿, No.4, pp.36-43 (1994).
- 2) 熊本忠彦, 伊藤 昭, 海老名毅: 支援対話におけるユーザ発話意図の認識, 信学論, Vol.J77-D-II, No.6, pp.1114-1123 (1994).
- 3) 伊藤 昭, 海老名毅, 熊本忠彦: 対話型計算機利用支援におけるユーザ質問の分類と支援回答戦略, 信学論, Vol.J77-D-II, No.7, pp.1319-1328 (1994).
- 4) 熊本忠彦, 伊藤 昭: 支援対話の解析に基づく対話処理方式の提案, 信学論, Vol.J77-D-II, No.8, pp.1492-1501 (1994).
- 5) Kumamoto, T., Ito, A. and Ebina, T.: Design and Construction of an Advisory Dialogue Database, *IEICE Trans. Information and Systems*, Vol.E78-D, No.4, pp.420-427 (1995).
- 6) 宇津呂武仁, 松本裕治: コーパスを用いた言語知識の獲得, 人工知能学会誌, Vol.10, No.2, pp.197-204 (1995).
- 7) 工藤育男, 井ノ上直己: コーパスに基づく共起知識の獲得とその応用, 人工知能学会誌, Vol.10, No.2, pp.205-212 (1995).
- 8) 北村美穂子, 松本裕治: 対訳コーパスを利用した翻訳規則の自動獲得, 情報処理学会論文誌, Vol.37, No.6, pp.1030-1040 (1996).
- 9) 松岡達雄, Hasson, R., Barlow, M., 古井貞照: テキストコーパスを用いた音声理解のための言語モデル自動獲得, 情報処理学会研究報告, Vol.95, No.120, pp.93-98 (1995).
- 10) 飯田 仁, 有田英一: 4階層プラン認識モデルを使った対話の理解, 情報処理学会論文誌, Vol.31, No.6, pp.810-821 (1990).
- 11) Nagata, M. and Morimoto, T.: An Information-Theoretic Model of Discourse for Next Utterance Type Prediction, 情報処理学会論文誌, Vol.35, No.6, pp.1050-1061 (1994).
- 12) Reithinger, N., Engel, R., Kipp, M. and Klesen, M.: Predicting Dialogue Acts for a Speech-To-Speech Translation System, *Proc. International Conference on Spoken Language Processing*, Philadelphia, USA, pp.654-657 (1996).
- 13) 黒橋禎夫, 木下恭子, 山田悟史, 長尾 真: 文タイプと文間関係の情報を付与したテキストコーパスの作成, 言語処理学会第4回年次大会発表論文集, pp.342-343 (1998).
- 14) 荒木雅弘, 市川 熹, 青柳達也, 石崎雅人, 伊藤敏彦, 柏岡秀紀, 熊谷智子, 小磯花絵, 田本真詞, 土屋 俊, 中里 収, 堀内靖雄, 前川喜久雄, 村上雄大, 山下洋一, 吉村 隆: 談話タグワーキンググループ活動報告, 人工知能学会研資, SIG-SLUD-9701-6, pp.31-36 (1997).
- 15) Jurafsky, D., Bates, R., Coccaro, N., Martin, R., Meteer, M., Ries, K., Shriberg, E., Stolcke, A., Taylor, P., and Van Ess-Dykema, C.: Automatic Detection of Discourse Structure for Speech Recognition and Understanding, *Proc. IEEE Workshop on Speech Recognition and Understanding*, Santa Barbara, USA (1997).
- 16) 滝永 登, 西本卓也, 新美康永: 観光案内対話を対象とした自動タグ付けの試み, 人工知能学会研資, SIG-SLUD-9801-2, pp.7-12 (1998).
- 17) 岩井康浩, 荒木雅弘, 堂下修司: 対話コーパスに対する意味・談話タグの推定手法, 人工知能学会研資, SIG-SLUD-9703-5, pp.27-32 (1998).
- 18) 駒谷和範, 荒木雅弘, 堂下修司: 発話内行為タグの推定法とそのタグ付与支援ツールへの組み込み,

人工知能学会研資, SIG-SLUD-9801-5, pp.29-34 (1998).

- 19) Peek, J.: *Mh and Xmh - E-mail for Users and Programmers*, O'Reilly & Associates (1991).
- 20) 益岡隆志, 田窪行則: 基礎日本語文法, くろしお出版, 東京 (1991).
- 21) 妙木 裕, 松本裕治, 長尾 真: 汎用日本語辞書および形態素解析システム, 第42回情報処理学会全国大会論文集, No.1C-9, pp.3-17-3-18 (1991).
- 22) Quinlan, J.R.: *C4.5: Programs for Machine Learning*, Morgan Kaufmann, San Mateo, CA (1993).
- 23) Weiss, S.M. and Kulikowski, C.A.: *Computer Systems That Learn: Classification and Prediction Methods from Statistics, Neural Nets, Machine Learning, and Expert Systems*, Morgan Kaufmann, San Mateo, CA (1991).

(平成 10 年 6 月 10 日受付)

(平成 11 年 3 月 5 日採録)



熊本 忠彦 (正会員)

昭和 41 年生。昭和 63 年筑波大学第三学群情報学類卒業。平成 2 年同大学院 (修士課程) 理工学研究科修了。同年, 郵政省通信総合研究所入所。現在, 同関西先端研究センター知識処理研究室主任研究官。平成 8 年筑波大学博士 (工学)。話し言葉による対話に興味を持ち, 頑健な自然言語対話システムの実現を目指している。電子情報通信学会, 人工知能学会, 言語処理学会各会員。



伊藤 昭 (正会員)

昭和 25 年生。昭和 47 年京都大学理学部物理学科卒業。昭和 54 年同大学院 (博士課程) 理学研究科修了。理学博士。同年, 郵政省電波研究所 (現通信総合研究所) 入所。以後, 通信ネットワーク, 知識処理, 対話システム, ヒューマンインタフェース, エージェントモデル, コミュニケーションの認知機構等の研究に従事。同関西先端研究センター知識処理研究室長, 研究調整官を経て, 平成 10 年 4 月より山形大学工学部電子情報工学科教授。電子情報通信学会, 人工知能学会, 日本ソフトウェア科学会, ACM, AAI 各会員。